Geospatial Data is Analytics Superfood

Jeff Jonas, IBM Fellow
Chief Scientist, Context Computing
http://www.twitter.com/jeffjonas
www.jeffjonas.typepad.com

© 2016 IBM Corporation

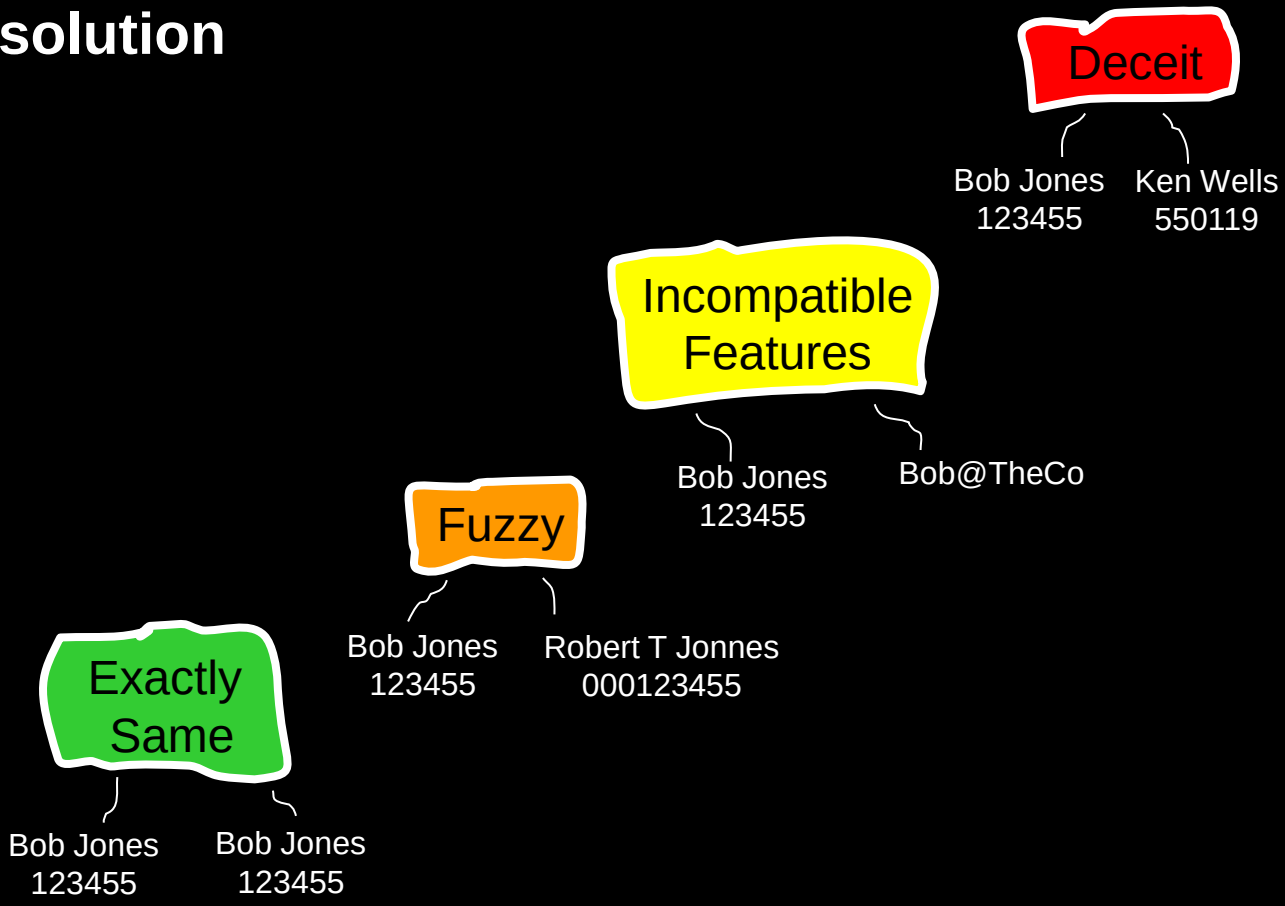
Jeff Jonas

IBM Fellow

Chief Scientist, Context Computing

- Founded Systems Research & Development (SRD) in 1985
- Architected, designed, developed roughly 100 systems over the last three decades
  - Defense, intelligence
  - Financial services
  - Gaming
  - Law enforcement
- Acquired by IBM in 2005
- Currently focused on Context Computing, Sensemaking and Privacy by Design (PbD)

Entity Resolution



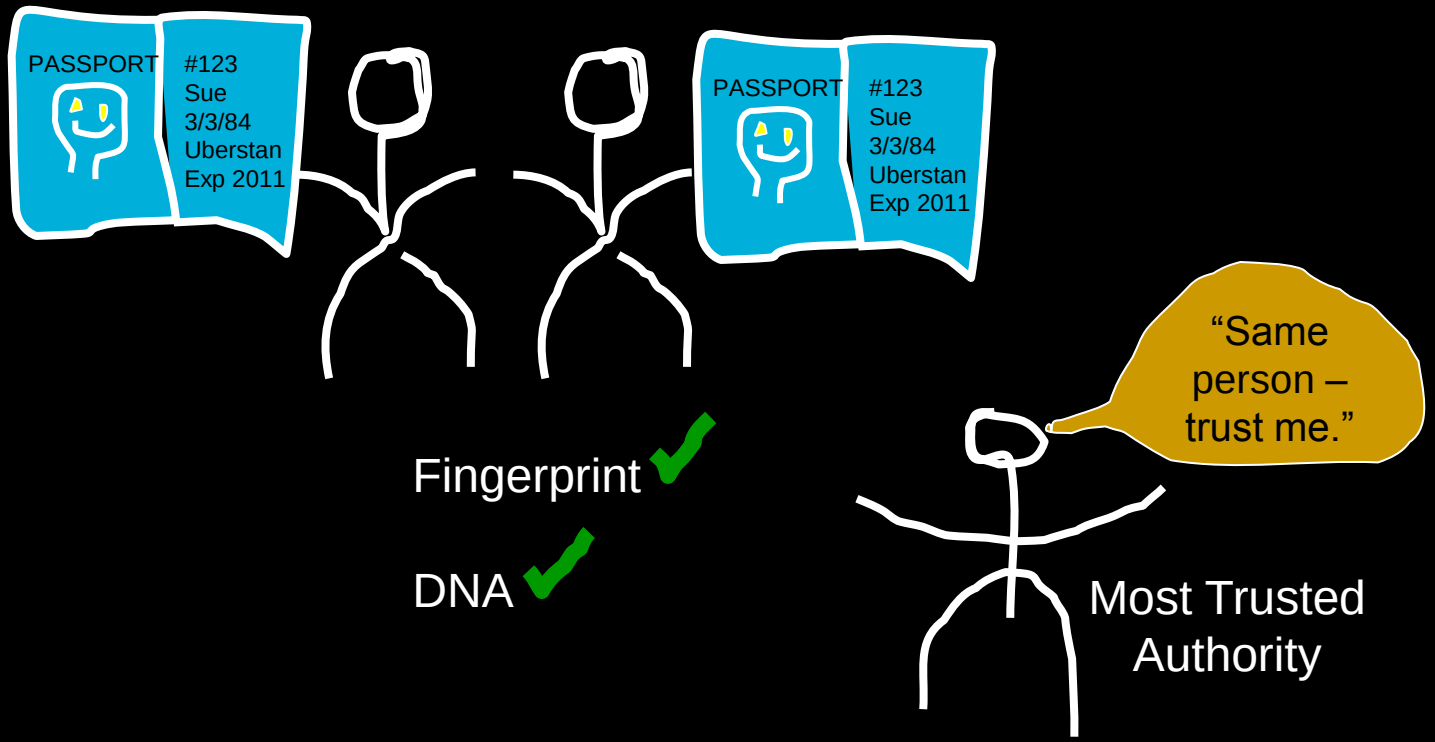
© 2016 IBM Corporation

Key Features Enable Entity Resolution

People	Cars	Router
Name	License Plate No.	Serial Number
Address	VIN	MAC Address
Date of Birth	Make	IP Address
Phone	Model	Make
Passport	Year	Model
Nationality	Color	Firmware Version
Biometric	Etc.	Etc.
Etc.		

© 2016 IBM Corporation

Consider Lying Identical Twins



© 2016 IBM Corporation

The same thing cannot be in two places … at the same time.

Two different things cannot occupy the same space … at the same time.

© 2016 IBM Corporation

Space & Time Enables Absolute Disambiguation

People	Cars	Router
Name	License Plate No.	Serial Number
When	When	When
Address	VIN	MAC Address
Where	Where	Where
Date of Birth	Make	IP Address
Phone	Model	Make
Passport	Year	Model
Nationality	Color	Firmware Version
Biometric	Etc.	Etc.
Etc.		

© 2016 IBM Corporation

"Life Arcs" Are Also Telling

Bill Smith
4/13/67
Salem, Oregon

Bill Smith
4/13/67
Seattle, Washington

Address History

Tampa, FL	2008-2015
Biloxi, MS	2005-2008
NY, NY	1996-2005
Tampa, FL	1984-1996

Address History

San Diego, CA	2005-2015
San Fran, CA	2005-2005
Phoenix, AZ	1990-2005
San Jose, CA	1982-1990

© 2016 IBM Corporation

© 2016 IBM Corporation

Space-Time-Travel

- Cell phones are generating a staggering amount of geo-locational data – 600B transactions per day being created in the US alone
- This data is being “de-identified” and shared with third parties – in volume and in real-time
- Your movement quickly reveals where you spend your time
- Re-identification (figuring out who is who) is somewhat trivial
- And, oh so powerful predictions …

© 2016 IBM Corporation

The 10 People I Spend the Most Time With (Not at Home and Not at Work)

- Michelle
- Renee
- Peggy
- Erin
- Joshua
- Ivan
- Bob
- Amanda
- Dane
- Wesley

© 2016 IBM Corporation

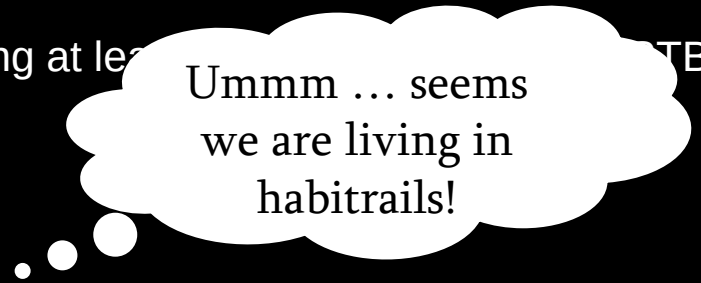
Unfair Advantage?

- The Uberstan intelligence service preempts the next mass protest in real-time
- A political opponent is crushed and resigns two days after announcing their candidacy

© 2016 IBM Corporation

Toying with Publically Available Cell Phone Data

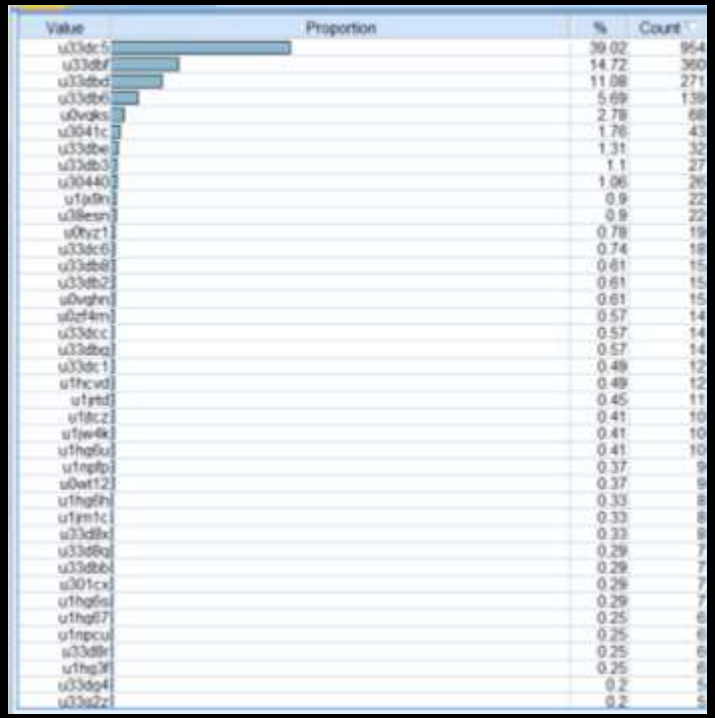
- 35,831 Call Data Records (CDRs)
  - 6 months: From 08-31-2009 through 02-27-2010
- 18,391 Total Number of Usable CDR's
  - Excluded CDRs with missing latitude, longitude, time, flow, or accuracy>250 meters
- 2,444 Hangouts
  - Minimum of 2 events, spanning at le TB
- The Pattern of Life
  - 130 Hangouts total
  - 64 Hangouts 3 or more times



© 2016 IBM Corporation

© 2016 IBM Corporation

Hangouts



Value	Proportion	%	Count
u33dc5		39.02	954
u33dbf		14.72	360
u33dbd		11.08	271
u33dbf		5.69	139
u0vqks		2.78	68
u3041c		1.76	43
u33dbe		1.31	32
u33db3		1.1	27
u30440		1.06	26
u1px9n		0.9	22
u38esn		0.9	22
u0fyz1		0.78	19
u33dc6		0.74	18
u33db8		0.61	15
u33db2		0.61	15
u0vghn		0.61	15
u0zf4m		0.57	14
u33dcc		0.57	14
u33dbq		0.57	14
u33dc1		0.49	12
u1fcwd		0.49	12
u1yrtd		0.45	11
u1jtcz		0.41	10
u1jw4k		0.41	10
u1hg6u		0.41	10
u1npfp		0.37	9
u0wt12		0.37	9
u1hg6h		0.33	8
u1jm1c		0.33	8
u33d8x		0.33	8
u33d8q		0.29	7
u33dbb		0.29	7
u301cx		0.29	7
u1hg6s		0.29	7
u1hg67		0.25	6
u1npcu		0.25	6
u33d8r		0.25	6
u1hg3f		0.25	6
u33dg4		0.2	5
u33d2z		0.2	5

© 2016 IBM Corporation

Getting to Know Malte Spitz

Six months of my life in 35,000 records
http://www.malte-spitz.de/blog/4103927.html

© 2016 IBM Corporation

Consequences

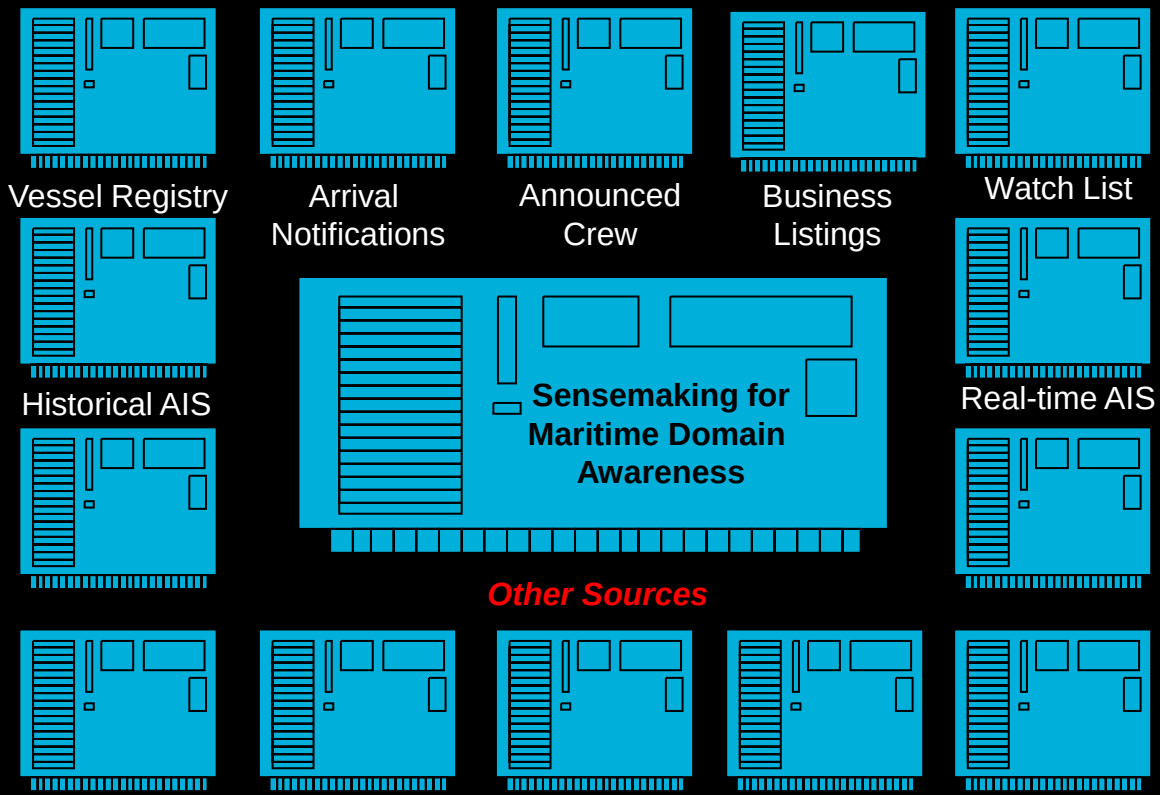
- Space-time-travel data is the ultimate biometric
- Adoption is now accelerating at a blistering pace
- It will unravel one’s secrets
- It will challenge existing notions of privacy
- It will enable enormous opportunity

© 2016 IBM Corporation

MARITIME DOMAIN AWARENESS

© 2016 IBM Corporation

Data Sources



© 2016 IBM Corporation

Vessels Hovering in Space-Time-Boxes

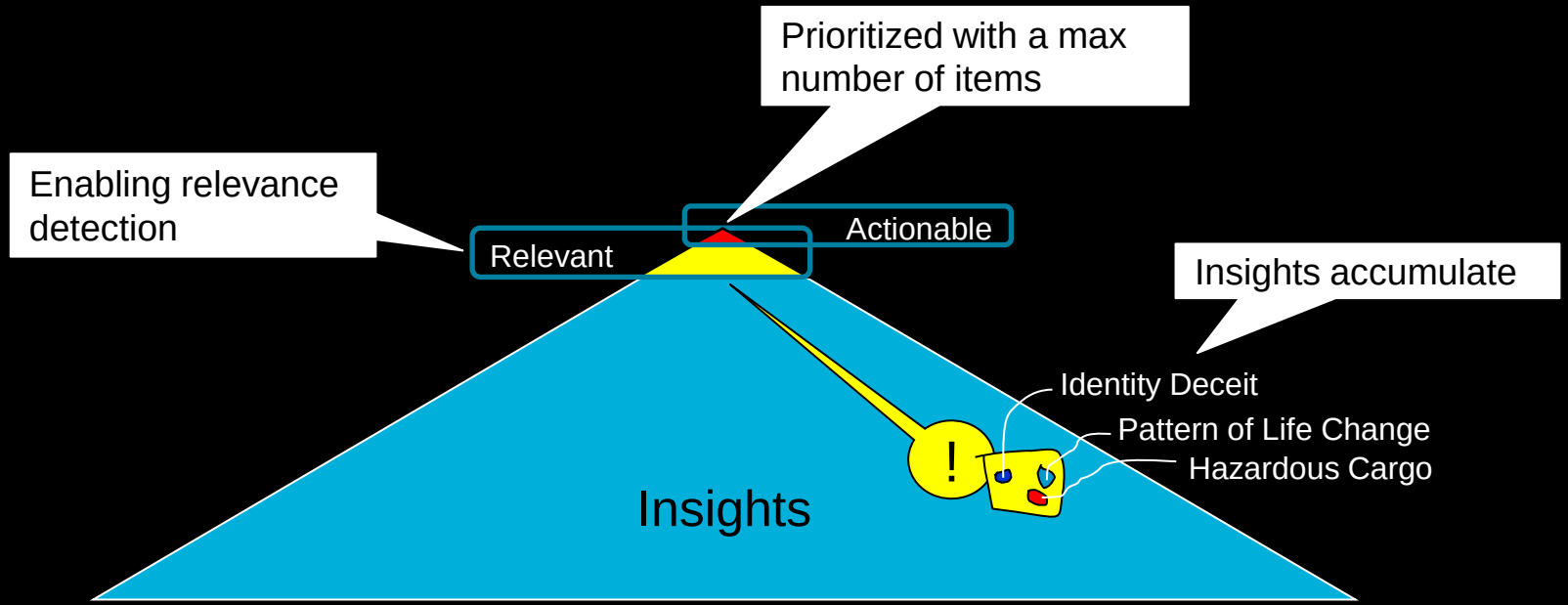
© 2016 IBM Corporation

Example Insights

Insight Code	Description	Score
WL_ENT	Watch Listed Entity	+85
WL_RELATIONS	Watch Listed Relationship	+50
IDENTITY_DECEIPT	Identity Deceit	+45
UNFAMILIAR	Unfamiliar Vessel	+30
CONFUSION	Confused Entity	+25
HAZ_CARGO	Hazardous Cargo	+20
POL_CHANGE	Pattern of Life Change	+10
FAMILIAR	Familiar Vessel	-20
TRUSTED_ENT	Trusted Entity	-40

© 2016 IBM Corporation

Developing Actionable Intelligence



© 2016 IBM Corporation

Sensing and Responding

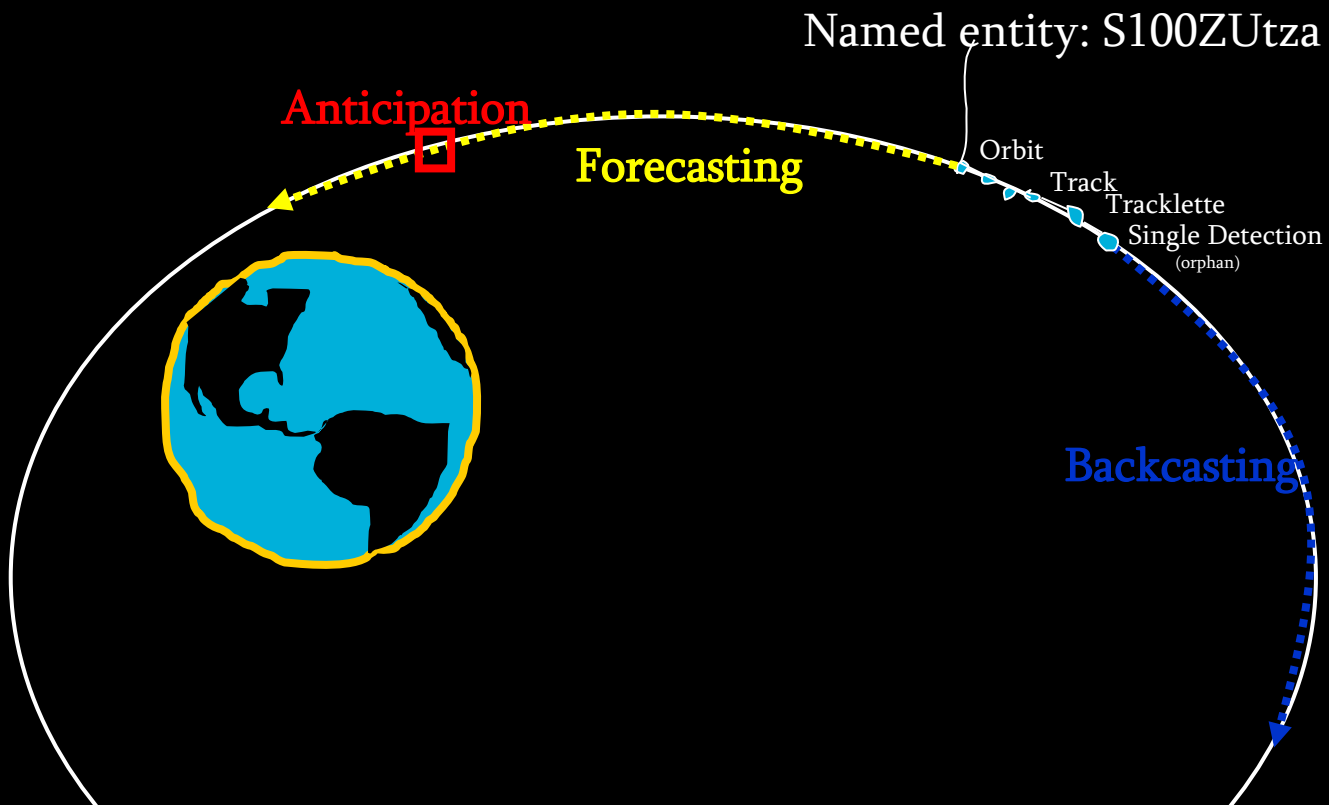
© 2016 IBM Corporation

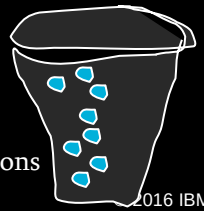
ASTEROID HUNTING

© 2016 IBM Corporation

Image courtesy of: Eva Lilly, Institute of Astronomy, University of Hawaii

From Orphans to Orbits





Single Detections
(trash)

2016 IBM Corporation

http://www.space.com/7854-slam-asteroids-suspected-space-collision.html

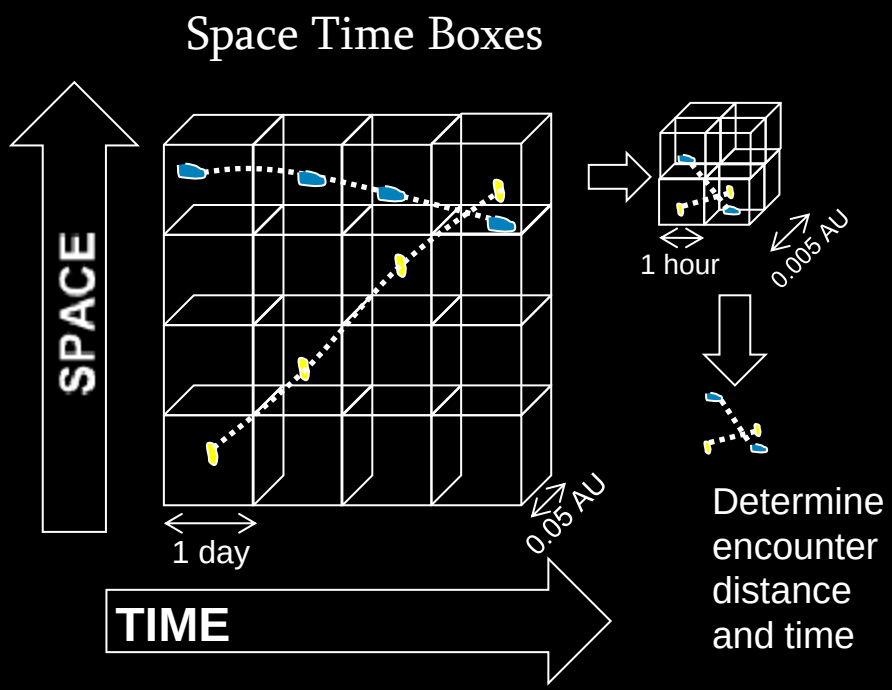
© 2016 IBM Corporation

"We have directly observed a collision between asteroids for the first time, instead of having to infer that they happened from million-year-old remains."

Colin Snodgrass
Planetary Scientist
Max Planck Institute for Solar System Research

© 2016 IBM Corporation

Detecting Colocation



© 2016 IBM Corporation

Computing 600k Asteroid Interactions over 25 Years

	Space-Time Box Method	N-body Simulation Method
Initial Analysis	2,880 CPU hours	10,000,000 CPU hours
Adding 1 New Trajectory	15 CPU minutes	4,000 CPU hours

4-5 orders of magnitude improvement

© 2016 IBM Corporation

June 12th, 2015

Hi Jeff & the gang,

I have great news! On Tuesday I happened to observe a close encounter you guys predicted - one 1 km and the other one 2 km in diameter!

To my knowledge this is the first case ever of direct observation of a close encounter in the small main belt asteroids. The closest point of encounter unfortunately happened during bright daylight in Hawaii, so I missed it …

Cheers!
Eva -

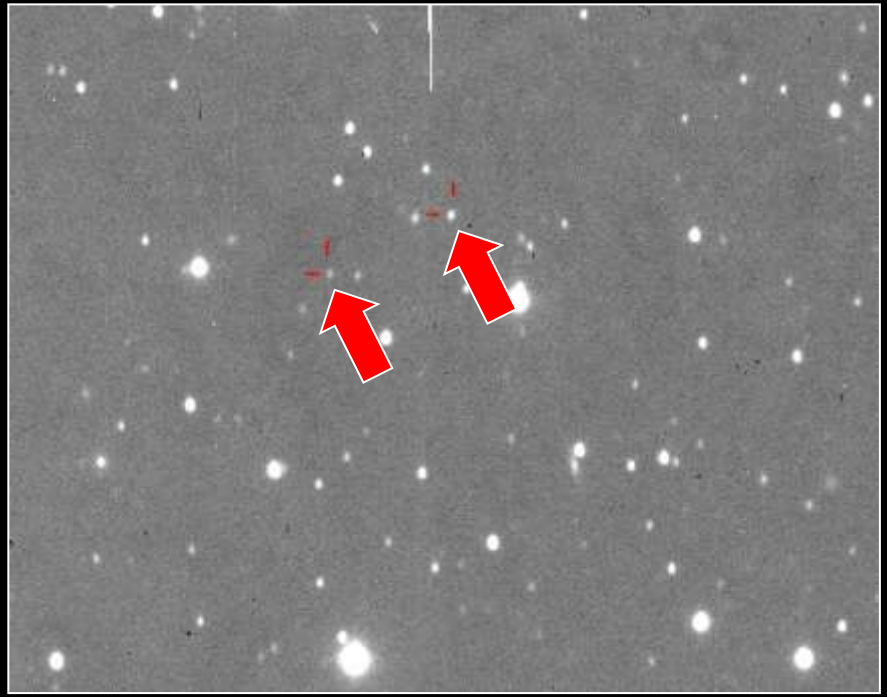


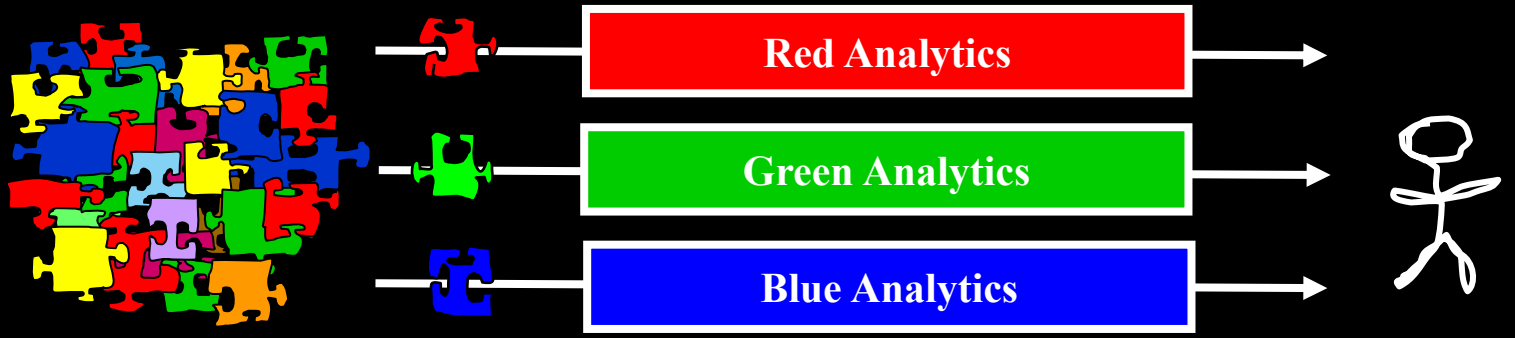
Image courtesy of Eva Lilly, Institute of Astronomy, University of Hawaii

© 2016 IBM Corporation

IN CLOSING

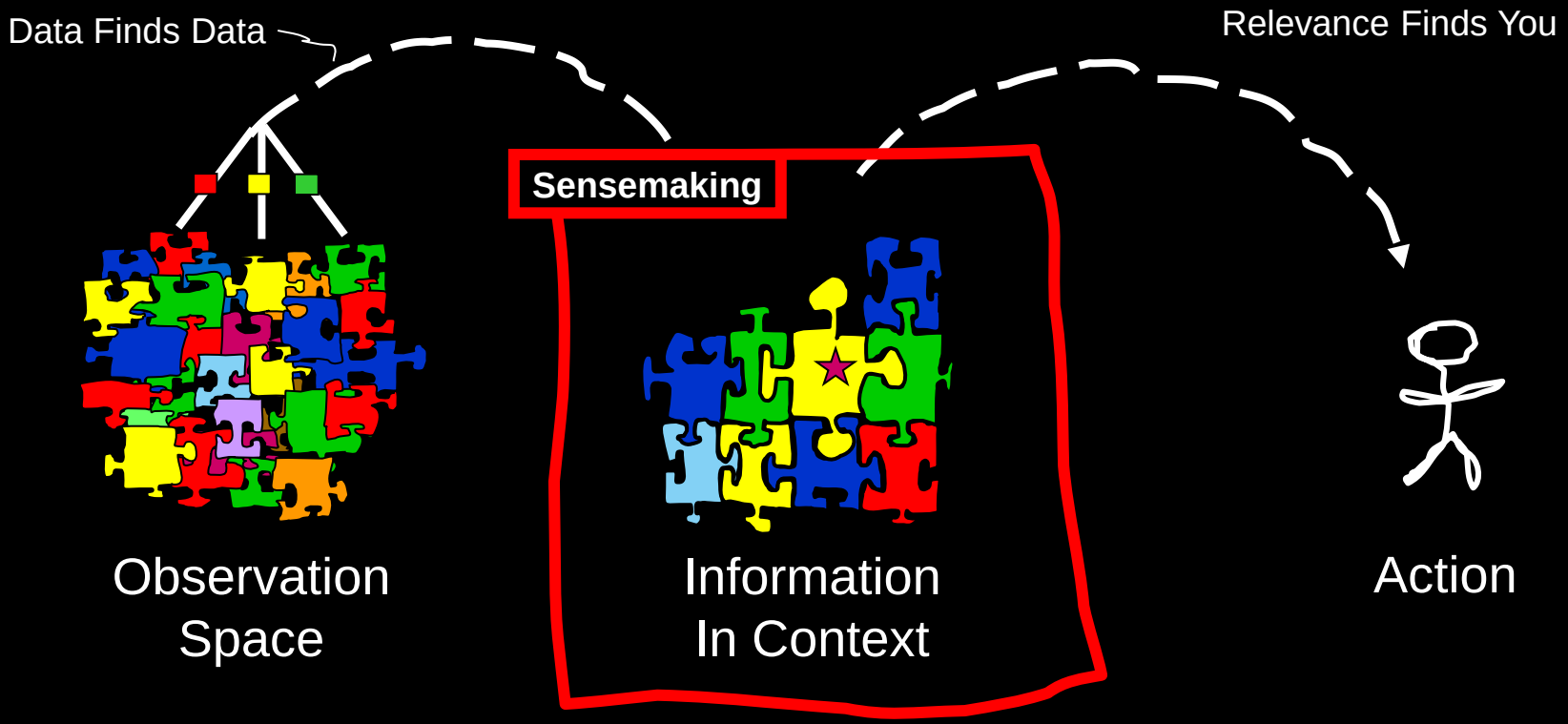
© 2016 IBM Corporation

Old School: Isolated Analytics



© 2016 IBM Corporation

Next: General Purpose Sensemaking



© 2016 IBM Corporation

The most competitive organizations

are going to make sense of what they are observing

fast enough to do something about it

while they are observing it.

© 2016 IBM Corporation

Geospatial Data is Analytics Superfood

Jeff Jonas, IBM Fellow
Chief Scientist, Context Computing
http://www.twitter.com/jeffjonas
www.jeffjonas.typepad.com

© 2016 IBM Corporation